

DOCUMENT RESUME

ED 030 848

AL 001 795

By-Shaffer, Stuart M.

The Measurement and Evaluation of Language Instruction.

Pub Date 7 Mar 69

Note- 13p.; Paper given at the Third Annual TESOL Convention, Chicago, Illinois, March 5-8, 1969.

EDRS Price MF-\$0.25 HC-\$0.75

Descriptors-Adjustment (to Environment), Auditory Discrimination, Diagnostic Tests, English (Second Language), Language Arts, Language Development, \*Language Instruction, \*Language Tests, Nonstandard Dialects, \*Pattern Drills (Language), \*Program Evaluation, Standard Spoken Usage, \*Ten1, Testing

Identifiers-Pattern Drills Program, Psychophysics

Developing test instruments for the Pattern Drills Program in the Pittsburg Public Schools has convinced the writer that the more test development activities and the teaching process reinforce each other, the stronger the program is. The Pattern Drills Program aims to develop bidialectism in non-standard English speakers by teaching standard English as a foreign language. The Drills reinforce and provide for "eventual automatic control of the standard pattern" by substitution practice. The contemporary psychophysics approach, described by Galanter in 1962 in terms of "detection," "recognition," "discrimination," and "scaling," can be used in testing for language development or for teaching language development. One reason for failure in teaching "correct standard English" is inappropriate measures. If a child cannot speak standard English at the appropriate time, we need to know whether it is because he cannot hear the difference, cannot mimic the difference, does not know the difference between different situations, or whether, although he has acquired all these "components," he just cannot combine them. Knowledge of this information would definitely have an effect on how we teach. (AMM)

## THE MEASUREMENT AND EVALUATION OF LANGUAGE INSTRUCTION

### Introduction

The past few years I have been concerned with the development of test instruments for the Pattern Drills Program in the Pittsburgh Public Schools. My attention to this aspect of the program has led me to the conclusion that the more test development activities and the teaching process reinforce each other, the stronger is the program.

One problem many teachers seem to have, is that they tend to see teaching as unreasonably complex. That is, teaching is viewed as a very complex set of interactions between the teacher and the student and the very complexity of these interactions makes the teaching situation less amenable to diagnosis and interruption. This is evidenced by the fact that many teachers tell me they feel that teaching is an art; that there are some people who are born to be teachers and there are some who are not. If this is true, it would be unreasonable to expect a school of education to teach students how to teach if in fact teaching itself is an inborn or indigenous phenomenon.

My conception of teaching is that teaching is a skill. The teacher-student interaction involves a set of communications which can be broken down and are amenable to measurement and evaluation. Further, communication between any individuals is subject to an analysis of its sub-components which would lead us to understand the total interaction. It is not difficult to see why teachers manifest some of the attitudes they do. It is partially a function of ego. It is true of everyone that we try to

AL 001 795

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE  
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

maximize the importance of our position so that we feel somewhat more important. (The implication is that importance and apparent complexity are related). This is something that everyone from the simplest laborer to the individual with the most complex occupation tries to do. Usually those with the more complex occupations tend to rely on the publicity generally afforded these occupations for displaying the complexity involved. An excellent example is the occupation of physician or perhaps psychologist.

I do feel that the major offender for giving teachers the wrong impression of their profession are schools of education. Regardless of the school of education a teacher has attended, he or she is graduated with a number of idealistic and unrealistic objectives. For example, teachers are taught to regard such things as a "happy child" to be important for learning. At no point does a school ever tell a teacher how to develop or recognize a happy child. Another problem is that the teacher is frequently told that he or she has to try to understand the background of the students, for it is only in this way that there will be adequate communication. Unfortunately, when the new teacher comes out of a school of education and goes into teaching, the first assignment in an urban school system is usually in one of the more disadvantaged schools and within that school, one of the more difficult classes. Immediately the problem becomes one of me versus them and all of the idealistic objectives and principles are soon forgotten.

### The Teaching of Language Skills

A teacher of language skills is indeed faced with a very difficult problem. The teacher has a clear idea of how the students should speak; that is, one can identify a good speaker and discriminate a good speaker from a bad speaker. But the classic procedures for teaching good speech have

always relied on teaching physical movements of the mouth. That is, where the tongue tip goes and similar instructions. It is only in recent years that we have become aware of the fact that students can put the tongue tip wherever they darn well please. Part of the problem is that they don't know when to put it there. The teacher who is faced with a student who speaks a very strong non-standard dialect frequently doesn't know where to begin. Further, if the student also uses an assortment of four letter words which was not part of the teacher's training, the teacher doesn't know how to modify the students' behavior so that these words are not used.

The field of behavior modification has introduced a number of concepts that are relevant for teachers. One such concept, shaping, refers to the process of rewarding those behaviors that are approximations of the ultimate behavior to be achieved. For example, if da student talks like dis, and not only do he talks like dis, but he also uses poor grammar, a teacher would be well advised to work on just one of these problems at a time, namely, enunciation or grammar, rather than try to work with all of the problems at once. Further, any learning that occurs on the part of the child must be reinforced or rewarded if a teacher wants this behavior to be maintained. If the teacher is working with individual components of the problem these components are then more available for reinforcement. If the teacher tries to reinforce a behavior which involves a large series of very complex skills, the teacher will frequently find that the child is receiving no reinforcement at all because these skills are developing differentially or at different rates. One behavior may be learned and forgotten before another is learned.

The Pittsburgh Board of Public Education is currently involved in a project known as the Pattern Drills Program. In our present society the advancement of the culturally deprived individual depends to a great extent on his quality of speech. It is becoming increasingly necessary for the individual to be able to express himself in the accepted speech of his particular region. This does not imply that he is discouraged from using his customary non-standard speech; with family and friends, the dialect which he is accustomed to is sufficient. In a situation such as a job interview however, this dialect may be unsuitable. The circumstances demand standard English, a more formal style of speech. The Pattern Drills Program was established for the purpose of equipping students with this faculty of bidialectism by teaching standard English as a foreign language.

The basic method of instruction for presenting the Pattern Drills to the students is the Pattern Drills themselves and three audio-visual aids. The Pattern Drills provide the actual instructional content for the program and assure that a particular pattern is correctly presented with respect to rhythm, continuity, and purity. For example, in a drill devoted to the standard use of "he doesn't," the students might repeat the following series of sentences after the teacher, each time focusing their attention on the changing direct object of the verb while the pattern the teacher wishes to reinforce remains constant and seemingly of secondary importance: he doesn't, he doesn't see the elephant, he doesn't see the giraffe, he doesn't see the tiger, he doesn't see the hippopotamus, etc. In order to reinforce and provide for eventual automatic control of the standard pattern, frequent substitution drills similar to the above example are presented in which students concentrate on non-essential substitutions in phrase or sentence content while they are repeating the desired pattern unchanged.



## The Testing of Language Skills

There is as yet no test devised which will accurately measure all aspects of the progress which students make in the Pattern Drills Program. There are a few informal test drills which the teacher includes in the curriculum. But these drills are only occasional checks and by no means evaluate the program as a whole. A testing instrument was needed to analyze the effectiveness of the entire program to determine the possible weak spots and to find out how far the student has advanced in his control of standard speech. The test must measure the actual speaking ability of the student as well as his ability to discern the appropriateness of standard and non-standard English.

It has been difficult to evaluate the measurement of speech objectively for raters may differ in their opinions on the correctness of pronunciation. According to Hitchman (1966), "There appear to be no records of test validation in the field of spoken English either in Britain or America.... The quality of the speech assessor is stressed rather than the actual rating scale." Various tests have been devised to evaluate articulation such as the Developmental Articulation Test, the Templin-Darley Tests of Articulation, and the Multiple Choice Intelligibility Test; but articulation is only one small aspect of the Pattern Drills Program.

There are very few tests in the field of listening. "In studying the neglect of listening, no where is there such yawning inadequacy as in the domain of standardized tests for measuring listening competence." (Hitchman, 1966). Not until 1959 did the Mental Measurements Yearbook list two listening tests (Brown-Carlson and STEP) and both of these were "wanting in many significant qualities." (Mental Measurements Yearbook, 1959).

At the George Peabody college demonstration school in Nashville, Tennessee, James W. Ney (1966) attempted to improve the writing ability of seventh grade students by the use of audio-lingual drills. His test consisted of a film containing no dialogue about which the students wrote an essay. The tests were scored by counting the structures which had been practiced extensively in class. This would not evaluate progress in the Pattern Drills Program, but it did show that this foreign language methodology can be used successfully "to improve composition as well as to strengthen both vocabulary and spelling." (Ney, 1966).

In Detroit Ruth Golden conducted a study in order to identify the oral language problems of culturally different students. This project was similar to the Pattern Drills Program. Dr. Golden's method of evaluation, however, was highly subjective since it was based primarily upon "impression and intuitive perceptions." None the less her study did indicate that improvement in speech habits, writing activities, and self-esteem are possible through the program. The District of Columbia (1965) introduced a successful language arts program in its public schools, "to develop the oral and written language facility and comprehension of culturally different children," their grades ranging from kindergarten through the third grade. The instrument for measuring proficiency in oral use of language was the Daily Language Facilities Test. The two scoring systems of the test are designed to measure: (1) the students' ability to use the language or dialect he learned at home and (2) the extent to which he speaks standard English. The students are required to respond verbally to a series of three pictures; their stories are then graded on a nine point scale. This test however is

designed for preschool children and does not measure the student's ability to recognize whether or not standard English would be applicable under certain conditions.

### Contemporary Psychophysics

Psychophysics refers to the measurement of physical phenomena through psychological procedures. For example, light can be measured with a photometer which will measure the actual intensity of the light or we can ask a subject to tell us how bright a light is on a scale from one to nine. A brightness measure would be considered a psychophysical measure whereas an intensity measure would be considered a physical measure. In contemporary psychophysics one breaks down the interaction of an individual with his physical environment in terms of a number of sequential variables. For example, our world is filled with many sounds; some of these sounds can be considered noise; some of these sounds can be considered actual signals. The relationship between the signal and the noise will determine to a great extent the intelligibility of the signal. In contemporary psychophysics we generally regard the first stage of interacting with a signal as being the Detection stage. That is, the individual must first be able to determine if he can hear anything at all that seems to be different from the background noise. The next stage is called the Recognition stage. After hearing or detecting the signal, does he recognize what the signal is? The first two stages can be exemplified by the questions "Is there anything there?" to represent the detection stage, and "What is it?" to represent the recognition stage.

The next stage in the interaction between an individual with a signal is called the Discrimination stage and here the question is,



"Is this different from that?" Basically we try to find out if the individual, even though he recognizes the kind of signal it is, can now give us more information about the signal. That is, is this a different signal from another signal? The fourth stage is called the Scaling stage and basically this stage asks the question, "How much of the signal is there?" These four stages are described by Galanter (1962).

An example of how we might view these stages in the study of language would be as follows: If we can imagine going to a foreign country where we have no familiarity with the language being spoken; the sounds would convey no meaning to us and hence might be considered noise. We would hear a large conglomerate of sounds but we would have no information about their meaning. If we can imagine being in a Polish speaking country and suddenly hearing a pattern of sounds other than Polish being spoken, we initially could tell it is a different language but we would not know which one. This can be conceived as the detection stage. As we are listening to the sounds and trying to decipher their meaning, we suddenly recognize that the pattern of sounds seems to be more familiar to us. This is considered the recognition stage. We now have somewhat more information about the signal we are hearing. If we listen more clearly or if we try to focus our attention on the source of this second language we can begin to discern individual words being spoken. That is, we can separate one word from another even though we might still not understand them. This state is called the discrimination stage. In the fourth stage, that is, the scaling stage, we would be able to actually comprehend what is being said. We can hear each word and we can comprehend the meaning of each word.

The same approach as is used in contemporary psychophysics can be used in developing a test for language development or for that matter even developing a sequence of steps for teaching language development. For example, one can see the sequence involved in the communication between one individual and another and more specifically, in the language expression between one person and another. Before we can teach a child when it is appropriate to use standard English and when it is appropriate to use non-standard English, we must first be confident that the child can hear the difference between the two forms of speech. If a child cannot hear this difference, then this child will never be able to learn to express the appropriate language form at the appropriate time. Our first test instrument would have to test the phenomenon of detection. That is, is there a perceptable difference between the two language forms? This question could be asked a number of ways. For example, one method would involve giving the child a standard sentence which might be "I ain't got none"; you then give the child a comparison sentence which might be "I don't have any," and another comparison sentence which might be "I ain't got none." The task would then be for the child to hear the three sentences and report which of the next two sentences agree with the first. In our example, of course, the second comparison is the same as the first. Notice that this is a relatively simple problem especially in this instance where we are using the same content. All of this information would be presented orally. A somewhat more difficult task but still within the area of detection would be the following: If we gave the child two language forms, such as "I ain't got none" and "dis is worser dan dat," we might ask the child whether these two forms are the same or different.

We could also present the child with, "Dis is worser dan dat" and "She helped her mother," and asked the child the same question. Are these two language forms the same or different? Naturally, we would expect this test to be somewhat more difficult than the first because there are fewer cues for the child to respond to. In the next stage of language development, we would ask the child to tell us if he knows when standard English is appropriate and when non-standard English is appropriate. There are a variety of ways of obtaining this kind of information. One way would be, for example, to present the child with a sentence such as "I ain't got none" and then asking the child if it would be appropriate to say this sentence to his mother, his friend on the corner, or his teacher. That is, we should be able to measure the student's awareness of the appropriate setting for using standard and non-standard dialect. The need for measuring the degree of this awareness is directly related to our desire to have students shift automatically from standard to non-standard speech and vice versa as the situation requires. A necessary first step to being able to select either standard or non-standard English to fit a given situation is for students to have an understanding of the milieu in which they find themselves. Attention needs to be given to determine if students can in fact appraise the situation in order to determine the propriety of one dialect or the other. This proposed instrument should make such a study possible. The third stage in the measurement of language development could very appropriately be called a mimicry stage. We are in effect asking the child if he can mimic standard English upon hearing it. For example, we might ask the child to repeat a sentence in standard English after we say it. This sentence

might be, "I was employed as a newspaper delivery boy." We would then ask the child to repeat this sentence after us. If he did this, we would be confident that the child could mimic standard English.

Instruments two and three are not necessarily sequential. We might desire to determine the child's mimicing capability before we decide to determine if the child can discriminate one situation from another. This probably would be equally valid.

For instrument number four we would expect the child to be able to combine the information that we know he has from instruments two and three. That is, knowing he has the individual components for speaking standard speech, now we want to know if this child can combine these components. This ability, while the most difficult to evaluate, is in the long run the most crucial of all the skills which the Pattern Drills Program seeks to develop. Until the present time, it has been very common for teachers to attempt to measure the child's ability to speak correct standard English after a certain period of time in the program. Most of the programs designed to accomplish this end have met with failure. One reason is that the measures have been inappropriate. If a child cannot speak appropriate standard English at the appropriate time, we do not know why this is the case: we do not know if the child cannot hear the difference, mimic the difference, whether or not he knows the difference between different situations, or whether although he has all the components he just cannot combine them. Knowledge of this information would definitely have an effect on how we teach this child. Rose Lee Nash (1967) wrote that "In the overall speech part of

the MES (More effective schools) program for the school 1965-1966 test results indicated that the second highest percentage of improvement was shown in developing audibility which was the most severe problem at the beginning of the school year. The second ranked problem, dialect, showed the third lowest percentage of improvement."

If we are attempting to make any kind of pitch at all here, it is that the teaching and measurement of English as a second language be conducted through the sequence of stages described herein. It is in this way that we will be able to ascertain the best method by which to teach a child to respond appropriately to his environment.



## REFERENCES

- Fifth Mental Measurements Yearbook, ed. Oscar Krisen Buros,  
The Gryphon Press, New Jersey, 1959.
- Galanter, Eugene. "Contemporary Psychophysics", ed. Brown,  
Roger, Eugene Galanter, Eckhard Hess, and George Mandler;  
New Directions in Psychology. Holt, Rinehart and Winston,  
New York, 1962.
- George Washington University Education Research Project.  
An Evaluation of the Language Arts Program of the District  
of Columbia, November, 1965.
- Hitchman, P. J. "The Validity and Reliability of Tests of Spoken  
English". The British Journal of Educational Psychology,  
36: 15-23, February, 1966.
- Nash, Rose Lee. "Teaching Speech Improvement to the Disadvantaged".  
Speech Teacher, 16C1, January, 1967 (69-73).
- Ney, James W. "Applied Linguistics in the Seventh Grade", English  
Journal, 55:895-897, October, 1966.